



Initial Semantics for Strengthened Signatures

André Hirschowitz, Marco Maggesi

► To cite this version:

André Hirschowitz, Marco Maggesi. Initial Semantics for Strengthened Signatures. 2012. hal-00762649

HAL Id: hal-00762649

<https://hal.science/hal-00762649>

Preprint submitted on 7 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Initial Semantics for Strengthened Signatures

André Hirschowitz Marco Maggesi

February 5, 2012

Contents

1	Introduction	2
1.1	The need for good notions of higher-order theory	2
1.2	Modules for modular higher-order syntax	2
1.3	Colimits of arities	3
1.4	Related and future work	3
1.5	Organization of the paper	4
2	Categories of modules over monads	4
2.1	Modules over monads	4
2.2	Morphisms of modules	5
2.3	The big category of modules	9
3	The category of arities	9
3.1	Arities	9
3.2	Morphisms of arities	10
3.3	Colimits of arities	11
3.4	Generalizations	11
4	Categories of representations	12
4.1	Signatures and their representations	12
4.2	The little category of representations	12
4.3	Representability	13
4.4	Modularity and the big category of representations	13
5	Strengthening signatures	14
5.1	Strengthened arities	14
5.2	Morphisms of strengthened arities	17
5.3	Representability of strengthened arities	18

6	Examples of strengthened syntax	22
6.1	Lambda-calculus modulo α -equivalence	22
6.2	Explicit composition operator	22
6.3	Syntax and semantics with explicit substitution	23

Abstract

We give a new general definition of arity, yielding the companion notions of signature and associated syntax. This setting is modular in the sense requested by [10]: merging two extensions of syntax corresponds to building an amalgamated sum. These signatures are too general in the sense that we are not able to prove the existence of an associated syntax in this general context. So we have to select arities and signatures for which there exists the desired initial monad. For this, we follow a track opened by Matthes and Uustalu [16]: we introduce a notion of strengthened arity and prove that the corresponding signatures have initial semantics (i.e. associated syntax). Our strengthened arities admit colimits, which allows the treatment of the λ -calculus with explicit substitution in the spirit of [10].

Contents

1 Introduction

1.1 The need for good notions of higher-order theory

Many programming or logical languages allow constructions which bind variables and this higher-order feature causes much trouble in the formulation, the understanding and the formalization of the theory of these languages. For instance, there is no universally accepted discipline for such formalizations: that is precisely why the POPLmark Challenge [4] offers benchmarks for testing old and new approaches. Although this problem may ultimately concern typed languages and their operational semantics, it already concerns untyped languages. In this work, we extend to new kinds of constructions our treatment of higher-order abstract syntax [12], based on modules and linearity.

1.2 Modules for modular higher-order syntax

First of all, we give a new general definition of arity, yielding the companion notion of signature. The notion is coined in such a way to induce a companion notion of representation of an arity (or of a signature) in a monad: such a representation is a morphism among modules over the given monad, so that an arity simply assigns two modules to each monad. There is a natural category of such representations of a signature and whenever it exists, the initial representation deserves the name of syntax associated with the

given signature. This approach enjoys modularity in the sense introduced by [10]: in our category of representations, merging two extensions of a syntax corresponds to building an amalgamated sum.

Our notion of arity (or signature) is too general in the sense that we are not able to build, for each signature, a corresponding initial representation. Following a track opened in Matthes-Uustalu [16], we define a fairly general notion of *strengthened* arity, yielding the corresponding notion of strengthened signature. Our main result (Theorem 5.15) says that any strengthened signature yields the desired initial representation. As usual, this initial object is built as a minimal fixpoint.

1.3 Colimits of arities

Understanding the syntax of the lambda-calculus with explicit substitution was already done in [10], where the arity for this construction was identified as a coend, hence a colimit, of elementary arities (see Section 6.3). Our main motivation for the present work (and for our next one) was to propose a general approach to syntax (and ultimately to semantics) accounting for this example in the spirit of our previous work [13]. This is achieved thanks to our second main result (Theorem 5.9) which states the existence of colimits in the category of (strengthened) arities.

ss:related

1.4 Related and future work

The idea that the notion of monad is suited for modeling substitution concerning syntax (and semantics) has been retained by many recent contributions on the subject (see e.g. [5, 10, 16]) although some other settings have been considered. Notably in [7] the authors work within a setting roughly based on operads (although they do not write this word down). The latter approach has been broadly extended, notably by M. Fiore [8]. Our main specificity here is the systematic use of the observation that the natural transformations we deal with are linear with respect to natural structures of module (a form of linearity had already been observed, in the operadic setting, see [9], section 4). Yet another approach is based on Lawvere Theories. This is clearly illustrated in the paper [14] where it is also outlined the link with the language of monads and put in an historical perspective.

The signatures we consider here are much more general than the signatures in [7], and cover the signatures appearing in [16, 10]. Note however that the latter works treat also non-wellfounded syntax, an aspect which we do not consider at all.

In our next work, we will propose a treatment of equational semantics for the present syntaxes. This approach should also be accommodated to deal with typed languages as done for elementary signatures in [17, 18, 2], or to model operational semantics as done for elementary signatures in [1].

1.5 Organization of the paper

Section 2 gives a succinct account about modules over a monad. Our new definitions of (higher-order) arity and signature are given in Section 3. Our solution to the problem of modularity appears in Section 4.3, while our strengthened arities appear in Section 5, together with the corresponding initiality theorem. Finally we give our examples in Section 6.

2 Categories of modules over monads

s:modules

2.1 Modules over monads

We recall only the definition and some basic facts about modules over a monad. See [13] for a more extensive introduction on this topic.

Let \mathbf{C} be a category. A monad over \mathbf{C} is a monoid in the category $\mathbf{C} \rightarrow \mathbf{C}$ of endofunctors of \mathbf{C} , i.e., a triple $R = (R, \mu, \eta)$ given by a functor $R: \mathbf{C} \rightarrow \mathbf{C}$, and two natural transformations $\mu: R^2 \rightarrow R$ and $\eta: I \rightarrow R$ such that the following equations hold:

$$\mu \cdot \mu R = \mu \cdot R\mu, \quad \mu \cdot \eta R = 1_R, \quad \mu \cdot R\eta = 1_R \quad (1) \quad \text{e:monad-axioms}$$

which are represented by the commutative diagrams

$$\begin{array}{ccc} R \cdot R \cdot R & \xrightarrow{\mu R} & R \cdot R \\ R\mu \downarrow & & \downarrow \mu \\ R \cdot R & \xrightarrow{\mu} & R \end{array} \quad \begin{array}{ccccc} I \cdot R & \xrightarrow{\eta R} & R \cdot R & \xleftarrow{R\eta} & R \cdot I \\ & \searrow 1_R & \downarrow \mu & \swarrow 1_R & \\ & & R & & \end{array}$$

Let R be a monad over \mathbf{C} .

Definition 2.1 (Modules). A left R -module is given by a functor $M: \mathbf{C} \rightarrow \mathbf{D}$ equipped with a natural transformation $\rho: M \cdot R \rightarrow M$, called *action*, which is compatible with the monad composition and identity:

$$\rho \cdot \rho R = \rho \cdot M\mu, \quad \rho \cdot M\eta = 1_M. \quad (2) \quad \text{e:module-axioms}$$

We will refer to the category \mathbf{D} as the *range* of M .

There is an obvious corresponding definition of right R -modules that we do not need to consider in this paper. From now on, we will write R -module instead of left R -module for brevity.

Let us show some trivial examples of modules.

Example 2.2. Every monad R is a module over itself, which we call the *tautological* module.

Example 2.3. For any functor $F: \mathbf{D} \rightarrow \mathbf{E}$ and any R -module $M: \mathbf{C} \rightarrow \mathbf{D}$, the composition $F \cdot M$ is an R -module (in the evident way).

Example 2.4. As an immediate consequence of the two above examples, the composition $R \cdot R$ is an R -module. This module will play a central role in our treatment of explicit substitution (see Section 6.2).

Example 2.5. For every object $W \in \mathbf{D}$ we denote by $\underline{W}: \mathbf{C} \rightarrow \mathbf{D}$ the constant functor $\underline{W} := X \mapsto W$. Then \underline{W} is trivially an R -module since $\underline{W} = \underline{W} \cdot R$.

Example 2.6. Let M_1, M_2 be two R -modules with the same range category \mathbf{D} . Assume that \mathbf{D} is a category with finite products. Then the product functor $M_1 \times M_2$ is an R -module (see Proposition 2.18 for a general statement).

For our purposes, one important example of module is given by the following general construction. Let \mathbf{C} be a category with finite colimits and a final object $*$.

Definition 2.7 (Derivation). For any R -module M with range \mathbf{D} , the *derivative* of M is the functor $M' := X \mapsto M(X + *)$. It is an R -module with the action $\rho': M' \cdot R \rightarrow M'$ defined the diagram

$$\begin{array}{ccc} M(R(X) + *) & \xrightarrow{\rho'_X} & M(X + *) \\ \downarrow M(R(i_X) + \eta_{X+*} \cdot *) & \nearrow \rho_{X+*} & \\ M(R(X + *)) & & \end{array} \quad (3) \quad \boxed{\text{e:action-derived}}$$

where $i_X: X \rightarrow X + *$ and $\eta: * \rightarrow X + *$ are the obvious maps. Derivation can be iterated, we denote by $M^{(k)}$ the k -th derivative of M .

Definition 2.8. Given a list of non negative integers $(a) = (a_1, \dots, a_n)$ we denote by $M^{(a)} = M^{(a_1, \dots, a_n)}$ the module $M^{(a_1)} \times \dots \times M^{(a_n)}$. Observe that, when $(a) = ()$ is the empty list, we have $M^{()} = *$ the final module.

2.2 Morphisms of modules

Definition 2.9 (Linearity). We say that a natural transformation of R -modules $\tau: M \rightarrow N$ is *linear* if it is compatible with actions:

$$\tau \cdot \rho^M = \rho^N \cdot \tau R.$$

We take linear natural transformations as morphisms among modules having the same range \mathbf{D} . It can be easily verified that we obtain in this way a category that we denote $\text{Mod}^{\mathbf{D}}(R)$.

This structure of category is for instance compatible with our product of modules, in the following sense.

Proposition 2.10. *If \mathbf{D} is a cartesian category, the product of modules is a cartesian product.*

This structure of category is also compatible with our derivation of modules, in the following sense.

Proposition 2.11. *Derivation yields an endofunctor of $\text{Mod}^{\mathcal{D}}(R)$. Moreover, if \mathcal{D} is a cartesian category, derivation is a cartesian endofunctor of $\text{Mod}^{\mathcal{D}}(R)$.*

In the case $\mathcal{C} = \mathcal{D} = \mathbf{Set}$ we have a natural *substitution morphism* $\sigma: M' \times R \longrightarrow M$. defined by the diagram

$$\begin{array}{ccc} M(X + *) \times R(X) & \xrightarrow{\sigma_X} & M(X) \\ w_X \downarrow & \nearrow \rho_X & \\ M(R(X)) & & \end{array}$$

where w is the map

$$w_X: (a, b) \mapsto M(\eta_X + \underline{b}), \quad \underline{b}: * \mapsto b$$

Lemma 2.12. *The transformation σ is linear.*

Proof. We have to prove the commutativity of the following diagram:

$$\begin{array}{ccc} (M' \times R) \cdot R & \xrightarrow{\sigma R} & M \cdot R \\ \rho' \times \mu \downarrow & & \downarrow \rho \\ M' \times R & \xrightarrow{\sigma} & M \end{array}$$

that is, given a set X and an element $(a, b) \in M(R(X) + *) \times R(R(X))$ we have to verify the identity

$$\rho_X(\sigma_{R(X)}(a, b)) = \sigma_X(\rho'_X(a), \mu_X(b)).$$

We reduce both sides of the equations. For the left hand side we have

$$2\rho_X(\sigma_{R(X)}(a, b)) = \rho_X(\rho_{R(X)}(M(\eta_{R(X)} + \underline{b})(a))) \quad (4) \quad \boxed{\text{e:sigma-lin-sx-1}}$$

$$= \rho_X(M(\mu_X \cdot \eta_{R(X)} + \mu_X \cdot \underline{b})(a)) \quad (5) \quad \boxed{\text{e:sigma-lin-sx-2}}$$

$$= \rho_X(M(1_{R(X)} + \mu_X \cdot \underline{b})(a)) \quad (6) \quad \boxed{\text{e:sigma-lin-sx-3}}$$

by using the definition of σ in (4), the associativity of ρ and the functoriality of M in (5) and the identity law for μ in (6). The right side is slightly more involved:

$$\sigma_X(\rho'_X(a), \mu_X(b)) = \rho_X(M(\eta_X + \underline{\mu_X(b)})(\rho'_X(a))) \quad (7) \quad \boxed{\text{e:sigma-lin-dx-1}}$$

$$= \rho_X\left(M(\eta_X + \mu_X \cdot \underline{b})(\rho_{X+*}(M(R(i_X) + \eta_{X+*} \cdot \underline{*})(a)))\right) \quad (8) \quad \boxed{\text{e:sigma-lin-dx-2}}$$

$$= \rho_X\left(\rho_{R(X)}(M(R(\eta_X + \mu_X \cdot \underline{b}) \cdot (R(i_X) + \eta_{X+*}(*)))(a))\right) \quad (9) \quad \boxed{\text{e:sigma-lin-dx-3}}$$

$$= \rho_X\left(M(\mu_X \cdot R(\eta_X + \mu_X \cdot \underline{b}) \cdot (R(i_X) + \eta_{X+*}(*)))(a)\right) \quad (10) \quad \boxed{\text{e:sigma-lin-dx-4}}$$

$$= \rho_X(M(1_{R(X)} + \mu_X \cdot \underline{b})(a)) \quad (11) \quad \boxed{\text{e:sigma-lin-dx-5}}$$

In (7) and (8) we unfold the definitions of σ and ρ' respectively. Next we use the naturality and the associativity of ρ respectively in (9) and (10), together with the functoriality of M . Finally (11) comes from the following auxiliary computation:

$$\begin{aligned} & \mu_X \cdot R(\eta_X + \mu_X \cdot R) \cdot (R(i_X) + \eta_{x+*} \cdot *) \\ &= \mu_X \cdot \left(R((\eta_X + \mu_X \cdot b) \cdot i_X) + R((\eta_X + \mu_X \cdot b) \cdot \eta_{x+*} \cdot *) \right) \end{aligned} \quad (12)$$

$$= \mu_X \cdot R(\eta_X) + \mu_X \cdot \eta_{R(X)} \cdot (\eta_X + \mu_X \cdot \underline{b}) \cdot * \quad (13)$$

$$= 1_{R(X)} + \mu_X \cdot \underline{b} \quad (14)$$

e:sigma-lin-dx-aux-1

e:sigma-lin-dx-aux-2

e:sigma-lin-dx-aux-3

where we use the functoriality of R in (12), the naturality of η in (13) and the unit properties (both) of μ in (14). \square

The substitution σ allows us to interpret the derivative M' as the “module M with one formal parameter added”. Higher-order derivatives have analogous morphisms (that we still denote with σ) $\sigma: M^{(b)} \times R^b \rightarrow M$. where $\sigma(t, m_1, \dots, m_b) \in M(X)$ is obtained by substituting $m_1, \dots, m_b \in R(X)$ in the successive stars of $t \in M^{(b)}(X) = M(X + * + \dots + *)$.

Example 2.13. The composition $\mu^R: R \cdot R \rightarrow R$ of a monad R is R -linear. The compatibility conditions are obtained by simply applying the functor R to the first and the third equations of (1).

x:RR->R

Example 2.14. The vertical composition of R -linear morphisms (when it makes sense) is R -linear. More explicitly, if $f_1: M_1 \rightarrow N_1$ is a morphism of $\text{Mod}^C(R)$ and $f_2: M_2 \rightarrow N_2$ is a morphism of $\text{Mod}^D(R)$ then $f_2 \cdot f_1: M_2 \cdot M_1 \rightarrow N_2 \cdot N_1$ which is R -linear.

x:M->MR-1

Example 2.15. Let M be any module. The natural transformation $M \cdot \eta_R = M(\eta_R): M \rightarrow M \cdot R$ is R -linear. Caveat: one would be tempted to deduce this linearity by the previous example by taking $f_1 = \eta_R$ and $f_2 = 1_M$, but this would not work because the functor I is not an R -module. Nevertheless the linearity of $M \cdot \eta_R$ holds and can be directly verified.

x:M->MR-2

Example 2.16. Let M be a module in $\text{Mod}^C(R)$. The natural transformation $\eta_R \cdot M: M \rightarrow R \cdot M$ is R -linear. The same remark as of the previous example applies.

x:M+*->M'

Example 2.17. Assume $C = \text{Set}$ and consider a module $M \in \text{Mod}^D(R)$. For each set X we have a natural map $M(X) + * \rightarrow M(X + *)$. This induces an R -linear morphism $M + * \rightarrow M'$.

Limits and colimits in the category of modules can be constructed pointwise:

p:lim-colim-lmod

Proposition 2.18 (Limits and colimits of modules). *If D is complete (resp. cocomplete), then $\text{Mod}^D(R)$ is complete (resp. cocomplete).*

In particular, we will often make use of the fact that, if the range category \mathbf{D} is cartesian, then the category $\text{Mod}^{\mathbf{D}}(R)$ is also cartesian.

We conclude this section by giving more considerations on the substitution morphism σ introduced earlier. In particular, we are concerned by the problem of stating a suitable associativity property for such substitution. To this end, we need to consider a slightly generalization and introduce an operator that we will call *operadic substitution* of the following kind:

$$\sigma_{p,(q_1,\dots,q_p)}: M^{(p)} \times R^{(q_1)} \times \dots \times R^{(q_p)} \longrightarrow M^{(q_1+\dots+q_p)}. \quad (15)$$

e:operadic_substitut

Such operator can be derived in general from the module action for suitable base category \mathbf{C} . Here for simplicity we give the construction in the case $\mathbf{C} = \text{Set}$. For every pair of indexes (i, j) with $1 \leq i \leq p$ and $1 \leq j \leq q_i$ let $*_{i,j}$ denote distinct singletons. Given a set X , define $X_i := X + \sum_{j=1}^{q_i} *_{i,j}$ and $X_0 = X + \sum_{i=1}^p \sum_{j=1}^{q_i} *_{i,j}$ and denote by $e_i: X_i \hookrightarrow X_0$ and $e_0: X \hookrightarrow X_0$ the associated natural inclusions. Thus we can identify $R(X_i)$ with $R^{(q_i)}(X)$ and $R(X_0)$ with $R^{(q_1+\dots+q_p)}(X)$. Analogously we consider a sequence of singletons $*_k$ for $1 \leq k \leq p$ and we identify $M(X + *_1 + \dots + *_p)$ with $M^{(p)}(X)$. Given an element $m_i \in R^{(q_i)}(X)$ we have an associated diagram

$$\begin{array}{ccc} *_{i,j} & \xrightarrow{\hat{m}_i} & R(X_i) \\ & \searrow \tilde{m}_i & \downarrow R(e_i) \\ & & R(X_0) \end{array}$$

where \hat{m}_i is the point map associated to m_i . Hence, to each tuple $(m_1, \dots, m_p) \in R^{(q_1)}(X) \times \dots \times R^{(q_p)}(X)$ we consider a the map $e = e_0 + \tilde{m}_1 + \dots + \tilde{m}_p: X + *_1 + \dots + *_p \longrightarrow R(X_0) = R^{(q_1+\dots+q_p)}(X)$. Such map e induces by the usual module substitution a map $M^{(p)}(X) \longrightarrow M^{(q_1+\dots+q_p)}(X)$ by the diagram

$$\begin{array}{ccc} M^{(p)}(X) & \xrightarrow{M(i)} & M(R^{(q_1+\dots+q_p)}(X)) \\ & \searrow s_{(m_1,\dots,m_p)} & \downarrow \rho_{X_0} \\ & & M^{(q_1+\dots+q_p)}(X) \end{array}$$

Then we define the *operadic substitution* $\sigma_{p,(q_1,\dots,q_p)}$ by

$$\sigma_{p,(q_1,\dots,q_p)}(t, m_1, \dots, m_p) := s_{(m_1,\dots,m_p)}(t).$$

Observe that our construction generalizes over any R -module M and thus for the monad R itself.

Our operadic substitution enjoys the usual operadic properties. Let us denote $\sigma_{p,(q_1,\dots,q_p)}(t, m_1, \dots, m_p)$ by simply $t \circ (m_1, \dots, m_p)$. Then we have

- Associativity: $m \circ (t_1 \circ (r_{1,1}, \dots, r_{1,k_1}), t_p \circ (r_{p,1}, \dots, r_{p,k_p})) = (m \circ (t_1, \dots, t_p)) \circ (r_{1,1}, \dots, r_{1,k_1}, \dots, r_{p,1}, \dots, r_{p,k_p})$.
- Identity: $m \circ (*, \dots, *) = m$.

2.3 The big category of modules

We already introduced the category $\text{Mod}^D(R)$ of modules with fixed base R and range D . It is often useful to consider a larger category which collects modules with different bases. To this end, we need first to introduce the notion of pull-back. In this section we will assume that all modules have the same range category D .

Definition 2.19 (Pull-back). Let $f: R \rightarrow S$ be a morphism of monads and M an S -module. The action $M \cdot R \xrightarrow{Mf} M \cdot S \xrightarrow{\rho} M$ defines an R -module which is called *pull-back* of M along f and noted f^*M . It can be easily verified that an S -linear natural transformation $g: M \rightarrow N$ is also an R -linear natural transformation $f^*g: f^*M \rightarrow f^*N$ and that $f^*: \text{Mod}^D(S) \rightarrow \text{Mod}^D(R)$ is a functor.

It can be easily verified that pull-back is well-behaved with respect to many important constructions. In particular:

Proposition 2.20. *Pull-back commutes with products and with derivation.*

Definition 2.21 (The big module category). We define the big module category BMod^D as follows:

- its objects are pairs (R, M) of a monad R and an R -module M .
- a morphism from (R, M) to (S, N) is a pair (f, m) where $f: R \rightarrow S$ is a morphism of monads, and $m: M \rightarrow f^*N$ is a morphism of R -modules. The category BMod^D comes equipped with a forgetful functor to the category of monads, given by the projection $(R, M) \mapsto R$.

3 The category of arities

In this section, we give our new notion of arity. The destiny of an arity is to have representations in monads. A representation of an arity a in a monad R should be a morphism between two modules $\text{dom}(a, R)$ and $\text{codom}(a, R)$. For instance, in the case of the arity a of a binary operation, we have $\text{dom}(a, R) := R^2$ and $\text{codom}(a, R) := R$. Hence an arity should consist of two halves, each of which assigns to each monad R a module over R in a functorial way. However, in all our natural examples, we have $\text{codom}(a, R) = R$ as above. Although this will no longer be the case in the typed case (which we do not consider here), we choose to restrict our attention to arities of this kind, where $\text{codom}(a, R)$ is R .

3.1 Arities

From now on we will consider only monads over the category Set and modules with range Set . For technical reasons, see Section 5, we restrict our

attention to the category of ω -cocontinuous endofunctors that we will denote $\text{End}^\omega(\mathbf{Set})$. Analogously we will write Mon^ω (resp. BMod^ω) for the full subcategory of monads (resp. of modules over these monads) which are ω -cocontinuous.

We recall that finite limits commute with filtered colimits in \mathbf{Set} . It follows that $\text{End}^\omega(\mathbf{Set})$ has finite limits and arbitrary (small) colimits. This is the key ingredient in the proofs of ω -cocontinuity for most of our functors.

Definition 3.1 (Arities). An *arity* is a right-inverse functor to the forgetful functor from the category BMod^ω to the category Mon^ω .

Now we give our basic examples of arities.

Example 3.2. The assignment $R \mapsto R$ is an arity which we denote by Θ .

Example 3.3. The assignment $R \mapsto *_R$, where $*_R$ denotes the final module over R is an arity which we denote by $*$.

Example 3.4. Given two arities a and b , the assignment $R \mapsto a(R) \times b(R)$ is an arity which we denote by $a \times b$. In particular $\Theta^2 = \Theta \times \Theta$ is the arity of any (first-order) binary operation and, in general Θ^n is the arity of n -ary operations.

Example 3.5. Given an arity a , for each non-negative integer n , the assignment $R \mapsto a(R)^{(n)}$ is an arity which we denote by $a^{(n)}$. As usual, we also set $a' := a^{(1)}$ and $a'' := a^{(2)}$.

Example 3.6. For each sequence of non-negative integers $s = (s_1, \dots, s_n)$, the assignment $R \mapsto R^{(s_1)} \times \dots \times R^{(s_n)}$ (see Definition 2.8) is an arity which we denote by $\Theta^{(s)}$. Arities of the form $\Theta^{(s)}$ are said *algebraic*. These algebraic arities are those which appear in [7].

Example 3.7. Given two arities a, b their composition $a \cdot b := R \mapsto a(R) \cdot b(R)$ is an arity.

3.2 Morphisms of arities

d:arity-morph

Definition 3.8. A morphism among two arities $a_1, a_2: \text{Mon}^\omega \rightarrow \text{BMod}^\omega$ is a natural transformation $m: a_1 \rightarrow a_2$ which, post-composed with the projection $\text{BMod}^\omega \rightarrow \text{Mon}^\omega$, becomes the identity.

We easily check that arities form a subcategory Ar of the category of functors from Mon^ω to BMod^ω .

Now we give some examples of morphisms of arities.

Example 3.9. The natural transformation $\mu: \Theta \cdot \Theta \rightarrow \Theta$ induced by the structural composition of monads is a morphism of arities (see Example 2.14).

Example 3.10. The two natural transformations $\Theta \cdot \eta$ and $\eta \cdot \Theta$ from Θ to $\Theta \cdot \Theta$ induced by the Examples 2.15 and 2.16 are morphisms of arities.

3.3 Colimits of arities

For the example of the λ -calculus with explicit substitution (see Section 6.3), we need some colimits of arities.

`t:arity-lim-colim`

Theorem 3.11. *The category of arities has finite limits and arbitrary (small) colimits.*

Proof. Limits and colimits of arities can be easily constructed point-wise in the larger functor category from ω -cocontinuous monads into arbitrary modules. Hence we must verify that these limit or colimit functors actually are arities. In particular, we have to prove that they produce modules that are ω -cocontinuous. This essentially amounts to show the compatibility of such limit or colimit with ω -colimits in the category of **Set**. Since arbitrary colimits always commute (in a cocomplete category), the case of colimits of arities poses no problem. More care is required for the case of limits, since limits and colimits do not commute in general. However, finite limits commute with filtered colimits (hence, in particular, ω -colimits) in the category of sets (see Theorem IX.2.1 in [15]). This allows us to construct finite limits in the category of arities. \square

3.4 Generalizations

Our notion of arity allows two natural extensions.

- As we already alluded above, we consider only arity for morphisms whose codomain is the monad, while we could consider arbitrary codomains. In the typed case, which we do not consider here, such arbitrary codomains will be the rule. In the present untyped context, one example that we can propose of a morphism whose codomain is not the monad is the operadic substitution introduced in (15) at the end of Section 2.2. Another relevant example is the construction $\text{app}_1 : \text{LC} \longrightarrow \text{LC}'$: this is, in the λ -calculus, the variant of the **app** construction which receives only one argument, and takes the fresh variable as second argument. One nice feature of this variant is the form it offers for the beta and eta rules:

$$\text{abs} \cdot \text{app}_1 = 1_{\text{LC}}, \quad \text{app}_1 \cdot \text{abs} = 1_{\text{LC}}'.$$

We plan to discuss in our next paper a formalism that allows to formulate these equations in our paper [11]. Note that morphisms of this kind, where the codomain is a derivative of the monad, can be “reduced” to the “raw” case where the codomain is the monad, and this is the reason why we do not consider this extension.

- The second possible extension builds upon the fact that new constructions create new modules, hence new arities. So given a base signature

Σ , a Σ -arity will be a pair of Σ -modules (see Definition 4.1 in the next section). Such a Σ -arity may be added to Σ , yielding a bigger signature. This picture allows for instance to consider partially defined constructions, like the predecessor. Here again we have no urgent example and this is the reason why we do not consider this extension.

4 Categories of representations

s:representations

4.1 Signatures and their representations

d:signatures

Definition 4.1 (Signatures). We define a signature $\Sigma = (O, \alpha)$ to be a family of arities $\alpha: O \rightarrow \text{Ar}$. A signature is said to be algebraic if it consists of algebraic arities.

d:representations

Definition 4.2 (Representation of an arity, of a signature). Given an ω -cocontinuous monad R over **Set**, we define a representation of the arity a in R to be a module morphism from $a(R)$ to R ; a representation of a signature Σ in R consists of a representation in R for each arity in Σ .

Example 4.3. The usual $\mathbf{app}: \mathbf{LC}^2 \rightarrow \mathbf{LC}$ is a representation of the arity Θ^2 into the monad \mathbf{LC} .

r:sig-eq-arity

Remark 4.4. Given a signature $\Sigma = (O, \alpha)$ we can consider the associated arity $a = \sum_{o \in O} \alpha_o$. Then it is easy to show that the category of Σ -representations is isomorphic to the category of a -representations.

In other words, in this particular setting we could avoid to mention signatures and reduce every construction and statement to mere arities. Nevertheless we maintain the distinction between the two notions in order to follow the usual terminology and, more importantly, in view of further developments where the analogous identification cannot be made (e.g., in the case of typed syntax).

4.2 The little category of representations

Definition 4.5. Given a signature $\Sigma = (O, \alpha)$, we build the category Mon^Σ of representations of Σ as follows. Its objects are ω -cocontinuous monads equipped with a representation of Σ . A morphism m from (M, r) to (N, s) is a morphism of monads from M to N compatible with the representations in the sense that, for each o in O , the following diagram of M -modules commutes:

$$\begin{array}{ccc} \alpha_o(M) & \xrightarrow{r_o} & M \\ a_o(m) \downarrow & & \downarrow m \\ m^*(\alpha_o(N)) & \xrightarrow{m^*s_o} & m^*N \end{array}$$

where the horizontal arrows come from the representations and the left vertical arrow comes from the functoriality of arities and $m: M \rightarrow m^*N$ is the morphism of monad seen as morphism of M -modules.

Proposition 4.6. *These morphisms, together with the obvious composition, turn Mon^Σ into a category which comes equipped with a forgetful functor to the category of monads.*

Proposition 4.7. *The category Mon^Σ has a final object.*

Proof. The final endofunctor $\underline{*}$, that is the constant functor that returns a singleton, has a trivial structure of monad and Σ -representation. It is immediate to verify that this is the final object of Mon^Σ . \square

4.3 Representability

We are primarily interested in the existence of an initial object in this category Mon^Σ .

Definition 4.8. A signature Σ is said representable if the category Mon^Σ has an initial object which we denote $\hat{\Sigma}$.

Theorem 4.9. *Algebraic signatures are representable.*

For more details we refer to our paper [12] (Theorems 1 and 2). We give below a more general result (Theorem 5.15).

4.4 Modularity and the big category of representations

It has been stressed in [10] that the standard approach (via algebras) to higher-order syntax lacks modularity. In the present section we show in which sense our approach via modules enjoys modularity. The key for this modularity is what we call the big category of representations.

Suppose that we have a signature $\Sigma = (O, a)$ and two subsignatures Σ_1 and Σ_2 covering Σ in the obvious sense, and let Σ_0 be the intersection of Σ_1 and Σ_2 . Suppose that these four signatures are representable (for instance because Σ is algebraic or strengthened in the sense defined below). Modularity would mean that the corresponding diagram of monads

$$\begin{array}{ccc} \hat{\Sigma}_0 & \longrightarrow & \hat{\Sigma}_1 \\ \downarrow & & \downarrow \\ \hat{\Sigma}_2 & \longrightarrow & \hat{\Sigma} \end{array}$$

is a pushout. The observation of [10] is that the diagram of raw monads is, in general, not a pushout. Since we do not want to change the monads, in order to claim for modularity, we will have to consider a category of enriched

monads. Here by enriched monad, we mean a monad equipped with some additional structure, namely a representation of some signature.

Our solution to this problem goes through the following “big” category of representations, which we denote by \mathbf{RMon} , where R may stand for representation or for rich:

- An object of \mathbf{RMon} is a triple (R, Σ, r) where R is a monad, Σ a signature, and r is a representation of Σ in R .
- A morphism in \mathbf{RMon} from $(R_1, (O_1, a_1), r_1)$ to $(R_2, (O_2, a_2), r_2)$ consists of a map $i := O_1 \rightarrow O_2$ compatible with a_1 and a_2 and a morphism m from (R_1, r_1) to $(R_2, i^*(r_2))$, where, for i injective, $i^*(r_2)$ should be understood as the restriction of the representation r_2 to the subsignature (O_1, a_1) .
- It is easily checked that the obvious composition turns \mathbf{RMon} into a category.

Now for each signature Σ , we have an obvious functor from \mathbf{Mon}^Σ to \mathbf{RMon} , through which we may see $\hat{\Sigma}$ as an object in \mathbf{RMon} . Furthermore, an injection $i: \Sigma_1 \rightarrow \Sigma_2$ obviously yields a morphism $i_* := \hat{\Sigma}_1 \rightarrow \hat{\Sigma}_2$ in \mathbf{RMon} . Hence our ‘pushout’ square of signatures as described above yields a square in \mathbf{RMon} . The proof of the following statement is straightforward.

Proposition 4.10. *Modularity holds in \mathbf{RMon} , in the sense that given a ‘cocartesian’ square of representable signatures as described above, the associated square in \mathbf{RMon} is cocartesian again.*

As usual, we will denote by \mathbf{RMon}^ω the full subcategory of \mathbf{RMon} constituted by ω -cocontinuous functors. It is easy to check that the previous statement is equally valid in \mathbf{RMon}^ω . Indeed, recall that, by our definition, the initial representation of representable signatures lies in \mathbf{RMon}^ω .

5 Strengthening signatures

`s:stren-arities`

Guided by the ideas of Matthes and Uustalu [16] we introduce in our framework the notion of *strengthened arity*.

5.1 Strengthened arities

For a category \mathbf{C} , let us denote by $\mathbf{End}_*^\omega(\mathbf{C})$ the category of ω -cocontinuous *pointed endofunctor*, i.e., the category of pairs (F, η) of an ω -cocontinuous endofunctor F of \mathbf{C} and a natural transformation $\eta: I \rightarrow F$ from the identity endofunctor to F . A morphism of pointed endofunctors $f: (F_1, \eta_1) \rightarrow (F_2, \eta_2)$ is a natural transformation $f: F_1 \rightarrow F_2$ satisfying $f \circ \eta_1 = \eta_2$.

Definition 5.1. A *strengthened arity* is a pair (H, θ) where H is an ω -cocontinuous endofunctor of $\mathbf{End}^\omega(\mathbf{Set})$ (i.e., $H \in \mathbf{End}^\omega(\mathbf{End}^\omega(\mathbf{Set}))$) and θ

is a natural transformation

$$\theta: H(-) \cdot \sim \longrightarrow H(- \cdot \sim)$$

(where $H(-) \cdot \sim$ and $H(- \cdot \sim)$ have to be understood as functors from $\text{End}^\omega(\text{Set}) \times \text{End}_*^\omega(\text{Set})$ to $\text{End}^\omega(\text{Set})$) satisfying

$$\theta_{X,(I,1_I)} = 1_{HX} \quad (16)$$

`e:theta_id`

and such that the following diagram is commutative

$$\begin{array}{ccc} H(X) \cdot Z_1 \cdot Z_2 & \xrightarrow{\theta_{X,(Z_1 \cdot Z_2, e_1 \cdot e_2)}} & H(X \cdot Z_1 \cdot Z_2) \\ & \searrow \theta_{X,(Z_1, e_1)} Z_2 \quad \nearrow \theta_{X \cdot Z_1, (Z_2, e_2)} & \\ & H(X \cdot Z_1) \cdot Z_2 & \end{array} \quad (17)$$

`e:theta_comp`

for every endofunctor X and pointed endofunctors $(Z_1, e_1), (Z_2, e_2)$.

We refer to θ as the *strength* of the arity.

Our first task is to make clear that our wording is consistent in the sense that a strengthened arity H somehow yields a genuine arity \tilde{H} . For this task, for each monad R we pose $\tilde{H}(R) := H(R)$ and we exhibit on it a structure of R -module. We do even slightly more by upgrading H into a *module transformer* in the following sense:

`d:mod-trans`

Definition 5.2. A module transformer is an endofunctor of the big module category BMod^ω which commutes with the structural forgetful functor $\text{BMod}^\omega \longrightarrow \text{Mon}^\omega$.

`p:H-is-mod-trans`

Proposition 5.3. Let (H, θ) be a strengthened arity. For every ω -cocontinuous monad R and ω -cocontinuous R -module M , we define the natural transformation $\rho^{H(M)}: H(M) \cdot R \longrightarrow H(M)$ as the composition $H(\rho^M) \cdot \theta_{M,R}$. Then $(H(M), \rho^{H(M)})$ is an R -module, and this construction upgrades H into a module transformer denoted by \hat{H} .

We call the restriction \tilde{H} of the module transformer \hat{H} to the category of monads the arity associated to the strengthened arity H .

Proof. Assume that R is a monad over Set and M an R -module. First we show the associativity property. Let us denote by τ the morphism $\tau = H(\rho^M) \cdot H(M\mu^R) \cdot \theta_{M,R \cdot R}$. We will show that the two triangles in the diagram

$$\begin{array}{ccc} H(M) \cdot R \cdot R & \xrightarrow{H(M)\mu^R} & H(M) \cdot R \\ \rho^{H(M)R} \downarrow & \searrow \tau & \downarrow \rho^{H(M)} \\ H(M) \cdot R & \xrightarrow{\rho^{H(M)}} & H(M) \end{array}$$

commute.

The commutativity of the upper triangle follows from the commutativity of the following diagram

$$\begin{array}{ccccc}
H(M) \cdot R \cdot R & \xrightarrow{H(M)\mu^R} & H(M) \cdot R & & \\
\theta_{M,R \cdot R} \downarrow & & \theta_{M,(R,\eta^R)} \downarrow & \searrow \rho^{H(M)} & \\
H(M \cdot R \cdot R) & \xrightarrow{H(M\mu^R)} & H(M \cdot R) & \xrightarrow{H(\rho^M)} & H(M)
\end{array}$$

which is ensured by the naturality of θ and by the very definition of $\rho^{H(M)}$.

Since M is a module, the morphism τ can also be expressed as $\tau = H(\rho^M) \cdot H(\rho^M R) \cdot \theta_{M,R \cdot R}$. Moreover, the morphism $\rho^M(M)R$ is given by the composition $H(\rho^M)R \cdot \theta_{M,R}R$, then the lower triangle of (5.1) unfolds as the following diagram

$$\begin{array}{ccccccc}
H(M) \cdot R \cdot R & \xrightarrow{\theta_{M,R \cdot R}} & H(M \cdot R) \cdot R & \xrightarrow{H(\rho^M)R} & H(M) \cdot R & & \\
& \searrow \theta_{M,R \cdot R} & \downarrow \theta_{M \cdot R,R} & & \downarrow \theta_{M,R} & \searrow \rho^{H(M)} & \\
& & H(M \cdot R \cdot R) & \xrightarrow{H(\rho^M tR)} & H(M \cdot R) & \xrightarrow{H(\rho^M)} & H(M)
\end{array}$$

which is commutative by the property (17) of θ , the naturality of θ and the definition of $\rho^{H(M)}$.

Finally, the identity property is given by the commutativity of the diagram

$$\begin{array}{ccccc}
H(M) \cdot I & \xrightarrow{\theta_{M,I}=1_{H(M)}} & H(M) & & \\
H(M)\eta^R \downarrow & & H(M\eta^R) \downarrow & \searrow H(1_M)=1_{H(M)} & \\
H(M) \cdot R & \xrightarrow{\theta_{M,R}} & H(M \cdot R) & \xrightarrow{H(\rho^M)} & H(M)
\end{array}$$

which follows by the property (16) of θ , the naturality of θ and the identity property of the module M . \square

Our next task is to upgrade our favorite examples of arities into strengthened arities.

x:Theta-stren

Example 5.4. The arity Θ comes from the strengthened arity (H, θ) where H and θ are the relevant identities.

x:final-stren

Example 5.5. The arity $*$ comes from the strengthened arity (H, θ) where H is the final endofunctor and θ is the relevant identity. This is the final strengthened arity.

x:ThetaTheta-stren

Example 5.6. The arity $\Theta \cdot \Theta$ comes from the strengthened arity (H, θ) where $H := X \mapsto X \cdot X$ and $\theta_{X,Y} : X \cdot X \cdot Y \longrightarrow X \cdot Y \cdot X \cdot Y := X \cdot \eta^Y \cdot X \cdot Y$; here we have written η^Y for the morphism from the identity functor to Y (remember that Y is pointed).

x:derivative-stren

Example 5.7. We will show in Proposition 5.10 that the derivation carries strengthened arities to strengthened arities.

r:stren-eq-arity

Remark 5.8. It is easy to prove that the category of strengthened arities is ω -cocomplete. Thus, as for the case of general arities and signatures (Remark 4.4), we can systematically reduce the study of strengthened signature to that of strengthened arities by replacing the given signature with the sum of its arities.

5.2 Morphisms of strengthened arities

Then we show how our basic constructions of arities upgrade into constructions in the category of strengthened arities, which we now describe. Its objects are strengthened arities and we take for morphisms from (H_1, θ_1) to (H_2, θ_2) those natural transformations $m: H_1 \rightarrow H_2$ which are compatible with θ_1 and θ_2 , that is, the diagram

$$\begin{array}{ccc} H_1(X) \cdot Z & \xrightarrow{\theta_1} & H_1(X \cdot Z) \\ m_X Z \downarrow & & \downarrow m_{X \cdot Z} \\ H_2(X) \cdot Z & \xrightarrow{\theta_2} & H_2(X \cdot Z) \end{array}$$

is commutative for every endofunctor X and every pointed endofunctor Z .

We start with limits and colimits.

t:stren-lim-colim

Theorem 5.9. *The category of strengthened arities has finite limits and arbitrary colimits.*

Proof. This is due to the four natural isomorphisms:

$$\begin{aligned} (\lim_i H_i)(F) \cdot G &\simeq \lim_i (H_i(F) \cdot G) , & (\text{colim}_i H_i)(F) \cdot G &\simeq \text{colim}_i (H_i(F) \cdot G) , \\ (\lim_i H_i)(F \cdot G) &\simeq \lim_i (H_i(F \cdot G)) , & (\text{colim}_i H_i)(F \cdot G) &\simeq \text{colim}_i (H_i(F \cdot G)) . \end{aligned}$$

The restriction on limits is due to the ω -cocontinuity condition. \square

Next, we take care of the derivation. We denote by D the endofunctor of **Set** given by $A \mapsto A + *$. For any other pointed endofunctor X over **Set** we have a natural transformation $w^X: D \cdot X \rightarrow X \cdot D$ given by

$$\begin{aligned} w_A^X &: X(A) + * \rightarrow X(A + *) \\ w_A^X &:= X(i_A) + \eta_{A+*} \cdot \underline{*} \end{aligned}$$

where $i_A: A \rightarrow A + *$ and $\underline{*}: * \rightarrow A + *$ are the inclusion maps.

p:stren-derived

Proposition 5.10. *If (H, θ) is a strengthened arity, then the pair (H', θ') , where $H' := X \mapsto H(X)'$ and $\theta'_{X,Z} := \theta_{X,Z} D \cdot H(X) w_Z$, is a strengthened arity. We call it the derivative of (H, θ) .*

Proof. The necessary verifications are straightforward. \square

Now we point out the possibility of composing strengthened arities (while there is no natural composition of arities).

d:stren-comp

Definition 5.11. If $H := (H, \rho)$ and $K := (K, \sigma)$ are two strengthened arities, their composition $H \cdot K$ is the pair $(H \cdot K, \theta)$ where θ is defined the following commutative diagram

$$\begin{array}{ccc} H(K(X)) \cdot Z & \xrightarrow{\theta_{X,(Z,e)}} & H(H(X \cdot Z)) \\ & \searrow \rho_{K(X),(Z,e)} \quad \nearrow H(\sigma_{X,(Z,e)}) & \\ & H(K(X) \cdot Z) & \end{array}$$

p:stren-comp

Proposition 5.12. *This composition turns strengthened arities into a strict monoidal category.*

Proof. The proof is a routine verification. \square

5.3 Representability of strengthened arities

Next, we turn to the main interest of strengthened arities (or signatures) which is that the fixed point we are interested in inherits a structure of monad.

p:stren-fixpoint

Proposition 5.13. *Let (H, θ) be a strengthened arity. Then the fixed point T of the functor $F := I + H$ is ω -cocontinuous and comes equipped with a structure of H -representation.*

Proof. (Compare with Theorem 10 in [16].)

For each natural number a we denote by T_a the functors $F^a(I)$, that is, $T_0 = I$ and $T_{a+1} = T_a + H(T_a)$, and we consider the associated natural inclusions $i_a: T_a \rightarrow T_{a+1}$, $r_a: H(T_a) \rightarrow T_{a+1}$, $i_{a,b}: T_a \rightarrow T_{a+b}$, $\eta_a: I \rightarrow T_a$. The colimit T of the diagram

$$T_0 \xrightarrow{i_0} T_1 \rightarrow \dots \rightarrow T_a \xrightarrow{i_a} T_{a+1} \rightarrow \dots$$

is a fixed-point for F . From the general fact that two colimits over independent indices commute, we easily deduce that T is ω -cocontinuous. Also observe that $H(T)$ is the colimit of $H(T_a)$, since H is ω -cocontinuous by definition. We endow the functor T of a structure of monad with the following constructions.

We define two families of natural transformations $\mu_{a,b}: T_a \cdot T_b \rightarrow T_{a+b}$ and $\rho_{a,b}: H(T_a) \cdot T_b \rightarrow H(T_{a+b})$. For $a = 0$ we simply take $\mu_{0,b} = 1_{T_b}$. Next we work by recursion on a and we define

$$\rho_{a,b}: H(T_a) \cdot T_b \xrightarrow{\theta_{T_a,T_b}} H(T_a \cdot T_b) \xrightarrow{H(\mu_{a,b})} H(T_{a+b})$$

and

$$\mu_{a+1,b}: T_a \cdot T_b + H(T_a) \cdot T_b \xrightarrow{\mu_{a,b} + \rho_{a,b}} T_{a+b} + H(T_{a+b}).$$

By taking the colimit over a and b , the families of natural transformations η_a , $\mu_{a,b}$, $\rho_{a,b}$ and r_a give us respectively a unit $\eta^T: I \rightarrow T$, a composition $\mu^T: T \cdot T \rightarrow T$, a T -action $\rho^T: H(T) \cdot T \rightarrow H(T)$ and a natural transformation $r^T: H(T) \rightarrow T$. It can be shown that μ^T , η^T , ρ^T and r^T verify the axioms of monad and H -representation. We omit most of the verifications which are routine. As a paradigmatic example, we show the associativity of the composition μ^T . The proof of the other properties are either substantially easier or recoverable with a similar argument *mutatis mutandis*.

Since T and μ^T are defined through a colimit construction from T_a and $\mu_{a,b}$, to prove the associativity of μ it suffices to show the commutativity of the diagram

$$\begin{array}{ccc} T_a \cdot T_b \cdot T_c & \xrightarrow{\mu_{a,b} T_c} & T_{a+b} \cdot T_c \\ T_a \mu_{b,c} \downarrow & & \downarrow \mu_{a+b,c} \\ T_a \cdot T_{b+c} & \xrightarrow{\mu_{a,b+c}} & T_{a+b+c} \end{array} \quad (18) \quad \boxed{\text{e:Tcomp_assoc}}$$

for all positive indices a , b and c .

We proceed by induction on a . The case $a = 0$ is immediate since $\mu_{0,b}: I \cdot T_b \rightarrow T_b$ is the identity. Next, assume that (18) be commutative for a given a . We have to prove the corresponding property for the case $a + 1$. Hence, by using the recursive equations for T_{a+1} and $\mu_{a+1,b}$, we have to prove the commutativity of the following diagram

$$\begin{array}{ccc} (T_a + H(T_a)) \cdot T_b \cdot T_c & \xrightarrow{\mu_{a,b} T_c + \rho_{a,b} T_c} & (T_{a+b} + H(T_{a+b})) \cdot T_c \\ T_a \mu_{b,c} + T_a \rho_{b,c} \downarrow & & \downarrow \mu_{a+b,c} + \rho_{a+b,c} \\ (T_a + H(T_a)) \cdot T_{b+c} & \xrightarrow{\mu_{a,b+c} + \rho_{a,b+c}} & T_{a+b+c} + H(T_{a+b+c}) \end{array} \quad (19) \quad \boxed{\text{e:Tcomp_ass_step}}$$

By distributing the compositions over the unions, our diagram can be split as the sum of two square diagrams, one of which is just the diagram (18), which commutes by induction hypothesis, and the other is the following

$$\begin{array}{ccc} H(T_a) \cdot T_b \cdot T_c & \xrightarrow{\rho_{a,b} T_c} & H(T_{a+b}) \cdot T_c \\ T_a \rho_{b,c} \downarrow & & \downarrow \rho_{a+b,c} \\ H(T_a) \cdot T_{b+c} & \xrightarrow{\rho_{a,b+c}} & H(T_{a+b+c}) \end{array} \quad (20) \quad \boxed{\text{e:Tcomp_ass_step'}}$$

After replacing all the occurrences of $\rho_{a,b}$ with its definition we obtain the

diagram

$$\begin{array}{ccccc}
& & H(T_a \cdot T_b) \cdot T_c & \xrightarrow{H(\mu_{a,b})T_c} & H(T_{a+b}) \cdot T_c \\
& \nearrow \theta_{T_a, (T_b, \eta_b)} T_c & \downarrow \theta_{T_a \cdot T_b, (T_c, \eta_c)} & & \downarrow \theta_{T_{a+b}, (T_c, \eta_c)} \\
H(T_a) \cdot T_b \cdot T_c & \xrightarrow{\theta_{T_a, (T_b \cdot T_c, \eta_b \eta_c)}} & H(T_a \cdot T_b \cdot T_c) & \xrightarrow{H(\mu_{a,b} T_c)} & H(T_{a+b} \cdot T_c) \\
\downarrow H(T_a) \mu_{b,c} & & \downarrow H(T_a \mu_{b,c}) & & \downarrow H(\mu_{a+b,c}) \\
H(T_a) \cdot T_{b+c} & \xrightarrow{\theta_{T_a, (T_{b+c}, \eta_{b+c})}} & H(T_a \cdot T_{b+c}) & \xrightarrow{H(\mu_{a,b+c})} & H(T_{a+b+c})
\end{array} \tag{21}$$

`e:Tcomp_ass_step''`

which is commutative thanks to the induction hypothesis and the properties of H and θ . \square

The previous proposition is the main step in the proof of the following

Lemma 5.14. *The fixed-point T of $F = I + H$ is the initial object in the category of the H -representations.*

Proof. Given any H -representation $r^W: H(W) \rightarrow W$, we have to show that there exists a unique morphism of representations $k: T \rightarrow W$. We will use the same notation and definition established in the previous proposition.

We start by defining a family of natural transformations $k_a: T_a \rightarrow W$ by recursion on a as follows. For $a = 0$ just take $k_0: T_0 = I \rightarrow W$ to be the unit of the monad W . Next, assuming that k_a has been defined, we pose $k_{a+1}: T_{a+1} = T_a + H(T_a) \rightarrow W$ as $k_{a+1} := k_a + r^W \cdot H(k_a)$. It follows at once that we have $k_{a+1} \cdot i_a = k_a$, hence this yields a natural transformation $k: T \rightarrow W$ as we wanted.

We have to check that k is a morphism of monad. It is immediate to verify that k respects the units of the two monads. Let us verify that k respects the composition. It suffices to prove that the following diagram commutes

$$\begin{array}{ccc}
T_a \cdot T_b & \xrightarrow{k_a k_b} & W \cdot W \\
\mu_{a,b} \downarrow & & \downarrow \mu^W \\
T_{a+b} & \xrightarrow{k_{a+b}} & W
\end{array} \tag{22}$$

`e:init-monad-morph`

since the desired commutativity follows by taking the colimit.

For $a = 0$ we have $k_0 = \eta^W$ and the commutativity of the previous diagram follows easily from the monad axioms of W . Then we proceed by induction on a . We first observe that, for all a and b the commutativity of

(22) entails the commutativity of the diagram

$$\begin{array}{ccccc}
H(T_a) \cdot T_b & \xrightarrow{H(k_a)k_b} & H(W) \cdot W & \xrightarrow{r^W} & W \cdot W \\
\rho_{a,b} \downarrow & & \downarrow \rho^W & & \downarrow \mu^W \\
H(T_{a+b}) & \xrightarrow{H(k_{a+b})} & H(W) & \xrightarrow{r^W} & W
\end{array} \tag{23}$$

e:init-mod-morph

since the right square comes from the representation axiom of W and the commutativity of the left square is given by

$$\begin{aligned}
\rho^W \cdot H(k_a)k_b &= H(\mu^W) \cdot \theta_{W, (W, \eta^W)} \cdot H(k_a)k_b \\
&= H(\mu^W) \cdot H(k_a k_b) \cdot \theta_{T_a, (T_b, \eta_b)} \\
&= H(k_{a+b}) \cdot H(\mu_{a,b}) \cdot \theta_{T_a, (T_b, \eta_b)} \\
&= H(k_{a+b}) \cdot \rho_{a,b}.
\end{aligned}$$

By taking the sum of (22) and (23) we get the commutative diagram

$$\begin{array}{ccc}
T_a \cdot T_b + H(T_a) \cdot T_b & \xrightarrow{k_a k_b + \rho^W \cdot H(k_a)k_b} & W \cdot W \\
\mu_{a,b} + \rho_{a,b} \downarrow & & \downarrow \mu^W \\
T_{a+b} + H(T_{a+b}) & \xrightarrow{k_{a+b} + r^W \cdot H(k_{a+b})} & W
\end{array}$$

which is precisely the diagram (22) for the indices $(a+1, b)$. This concludes the proof that k is a morphism of monads.

Next we have to verify that k is a morphism of representations. Since H is a module transformer (Proposition 5.3) we already know that $H(k): H(T) \rightarrow H(W)$ is linear. Moreover, from the definition of k_a follows at once that we have $k_{a+1} \cdot r_a = r^W \cdot H(k_a)$. By taking the colimit over a we get that k commutes with the representation morphisms.

Finally we have to prove that k is unique. Then suppose that $h: T \rightarrow W$ be an arbitrary morphism of representations. It suffice to show the commutativity of the following diagram

$$\begin{array}{ccc}
T & \xrightarrow{h} & W \\
\uparrow u_a & \nearrow k_a & \\
T_a & &
\end{array} \tag{24}$$

e:init-unique

where the natural transformations $u_a: T_a \rightarrow T$ are the structural morphisms of the colimit T . Once more, we proceed by induction on a . The case $a = 0$ reduces to the compatibility of the monad morphism h with the unit of the two monads T and W . Assuming that (24) commutes for a given a , we can immediately deduce the commutativity of the left diagram in (25)

and thus the commutativity of the right diagram in (25) which is diagram (24) for the case $a + 1$.

$$\begin{array}{ccc}
 T & \xrightarrow{h} & W \\
 \uparrow r^T & & \uparrow r^W \\
 H(T) & \xrightarrow{H(h)} & H(W) \\
 \uparrow H(u_a) & \nearrow H(k_a) & \\
 H(T_a) & &
 \end{array}
 \qquad
 \begin{array}{ccc}
 T & \xrightarrow{h} & W \\
 \uparrow u_a + r^T \cdot H(u_a) & \nearrow k_a + r^W \cdot H(k_a) & \\
 T_a + H(T_a) & &
 \end{array}
 \quad (25)$$

`e:init-unique_aux`

□

The previous lemma leads immediately (Remark 5.8) to the following result.

`t:initsynt`

Theorem 5.15. *Strengthened signatures are representable.*

6 Examples of strengthened syntax

`s:examples`

6.1 Lambda-calculus modulo α -equivalence

One paradigmatic example of syntax with binding is the λ -calculus. We denote by $\Lambda(X)$ the set of lambda-terms up to α -equivalence with free variables ‘indexed’ by the set X . It is well-known [6, 3, 12] that Λ has a natural structure of cocontinuous monad where the monad composition is given by variable substitution.

It can be easily verified that application and abstraction are Λ -linear natural transformations $\mathbf{app}: \Lambda^2 \rightarrow \Lambda$ and $\mathbf{abs}: \Lambda' \rightarrow \Lambda$. That is, Λ is a monad endowed with a representation ρ of the signature $\Sigma = \{\mathbf{app}: \Theta^2, \mathbf{abs}: \Theta'\}$.

The monad Λ is initial in the category \mathbf{Mon}^Σ of ω -cocontinuous monads equipped with a representation of the signature Σ .

This is an example of algebraic signature and thus already treated by other previous works [12, 13, 7]. Here we simply remark that our new theory covers such a classical case.

6.2 Explicit composition operator

`ss:ex-expl-comp`

We now consider our first example of non-algebraic signature. On any monad R , we have the composition operator (also called join operator)

$$\mu^R: R \cdot R \rightarrow R$$

which has arity $\Theta \cdot \Theta$. We will refer to the μ^R operator as the *implicit* composition operator. An interesting problem is to see if this kind of operators

admits a corresponding *explicit* version, i.e., if they can be implemented as a syntactic construction.

As we have seen in Example 5.6 $\Theta \cdot \Theta$ is a strengthened arity hence we can build syntaxes with explicit composition operator of kind

$$\text{join}: \Theta \cdot \Theta \longrightarrow \Theta.$$

Of course, this is only a *syntactic* composition operator, in the sense that it does not enjoy several desirable conversion rules like associativity, two-side identity and the obvious compatibility rules with the other syntactic constructions present in the signature. In our next work we will show how to construct such kind of *semantic* composition operator.

Let us mention that given a monad R , the unit $\eta_R: I \longrightarrow R$ is not an R -linear morphism (in fact, I is not even an R -module in general). For this reason we cannot treat examples of syntax with explicit unit.

6.3 Syntax and semantics with explicit substitution

ss:ex-expl-subst

On any monad R , we have a series of substitution operators

$$\sigma_n: R^{(n)} \cdot R^n \longrightarrow R$$

which simultaneously replace n formal arguments in a term with n given terms.

As for the case of the composition of the previous section, we can easily construct examples of syntaxes with explicit substitution constructions $\text{subst}_n: \Theta^{(n)} \times \Theta^n \longrightarrow \Theta$ and add a set of equations that can be easily devised and that will eventually impose $\text{subst}_n = \sigma_n$. This is easily done on the track given in the previous section and we avoid to give the details.

Instead we want to focus to a slightly different point. As observed by Ghani and Uustalu [10], these substitution morphisms satisfy a series of compatibility relations which mean that they come from a single morphism $\text{subst}: C \longrightarrow \Theta$ where C is identified as the coend

$$C = \int^{A:\text{Fin}} \Theta^{(A)} \times \Theta^A.$$

Here Fin stands for the category of finite sets, Θ^A denotes the cartesian power and $\Theta^{(A)}$ is defined by $\Theta^{(A)}(R, X) := R(X + A)$. Since coends are special colimits [15], and strengthened arities admit colimits, we just have to check that the bifunctorial arity $(A, B) \mapsto \Theta^{(A)} \times \Theta^B$ factors through the category of strengthened arities. As far as objects are concerned, this follows from our results in Section 5. The verification of the compatibility of the corresponding “renaming” and “projection” morphisms with the strengthened structures is straightforward.

References

- [1] Benedikt Ahrens (2011): *Modules over relative monads for syntax and semantics*. ArXiv e-prints [arXiv:1107.5252](https://arxiv.org/abs/1107.5252).
- [2] Benedikt Ahrens & Julianna Zsidó (2011): *Initial Semantics for higher-order typed syntax in Coq*. *Journal of Formalized Reasoning* 4(1), pp. 25–69.
- [3] Thorsten Altenkirch & Bernhard Reus (1999): *Monadic Presentations of Lambda Terms Using Generalized Inductive Types*. In: *CSL*, pp. 453–468.
- [4] B. Aydemir, A. Bohannon, M. Fairbairn, J. Foster, B. Pierce, P. Sewell, D. Vytiniotis, G. Washburn, S. Weirich & S. Zdancewic (2005): *Mechanized metatheory for the masses: The POPLmark Challenge*. In: *Proceedings of the Eighteenth International Conference on Theorem Proving in Higher Order Logics (TPHOLs 2005)*.
- [5] Richard Bird & Ross Paterson (1999): *Generalised Folds for Nested Datatypes*. *Formal Aspects of Computing* 11(2), pp. 200–222.
- [6] Richard S. Bird & Ross Paterson (1999): *De Bruijn Notation as a Nested Datatype*. *Journal of Functional Programming* 9(1), pp. 77–91.
- [7] Marcelo Fiore, Gordon Plotkin & Daniele Turi (1999): *Abstract Syntax and Variable Binding*. In: *LICS '99: Proceedings of the 14th Annual IEEE Symposium on Logic in Computer Science*, IEEE Computer Society, Washington, DC, USA, p. 193.
- [8] Marcelo P. Fiore (2008): *Second-Order and Dependently-Sorted Abstract Syntax*. In: *LICS*, IEEE Computer Society, pp. 57–68. Available at <http://doi.ieeecomputersociety.org/10.1109/LICS.2008.38>.
- [9] Marcelo P. Fiore & Daniele Turi (2001): *Semantics of Name and Value Passing*. In: *Logic in Computer Science*, pp. 93–104.
- [10] Neil Ghani, Tarmo Uustalu & Makoto Hamana (2006): *Explicit substitutions and higher-order syntax*. *Higher-order and Symbolic Computation* 19(2–3), pp. 263–282.
- [11] André Hirschowitz & Marco Maggesi: *Initial Algebraic Semantics with Equations*. Coming soon.
- [12] André Hirschowitz & Marco Maggesi (2007): *Modules over Monads and Linearity*. In Daniel Leivant & Ruy J. G. B. de Queiroz, editors: *WoLLIC, Lecture Notes in Computer Science* 4576, Springer, pp. 218–237. Available at http://dx.doi.org/10.1007/978-3-540-73445-1_16.

- [13] André Hirschowitz & Marco Maggesi (2010): *Modules over monads and initial semantics*. *Information and Computation* 208(5), pp. 545–564. Special Issue: 14th Workshop on Logic, Language, Information and Computation (WoLLIC 2007).
- [14] Martin Hyland & John Power (2007): *The category theoretic understanding of universal algebra: Lawvere theories and monads*. *Electronic Notes in Theoretical Computer Science* 172, pp. 437–458.
- [15] Saunders Mac Lane (1998): *Categories for the working mathematician*, second edition. *Graduate Texts in Mathematics* 5, Springer-Verlag, New York.
- [16] Ralph Matthes & Tarmo Uustalu (2004): *Substitution in non-wellfounded syntax with variable binding*. *Theor. Comput. Sci.* 327(1-2), pp. 155–174, doi:<http://dx.doi.org/10.1016/j.tcs.2004.07.025>.
- [17] Julianna Zsidó (2005/06): *Le lambda calcul vu comme monade initiale*. Master’s thesis, Université de Nice – Laboratoire J. A. Dieudonné. Mémoire de Recherche – master 2.
- [18] Julianna Zsidó (2010): *Typed Abstract Syntax*. Ph.D. thesis, University of Nice, France. <http://tel.archives-ouvertes.fr/tel-00535944/>.